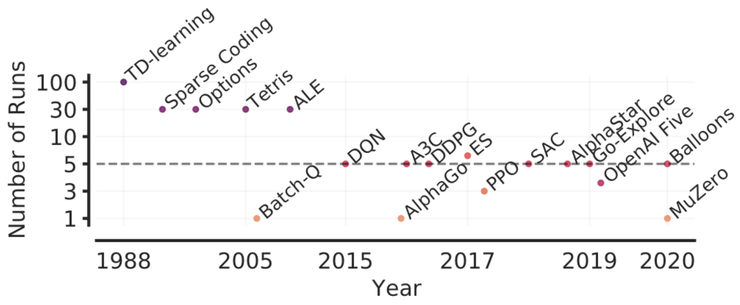
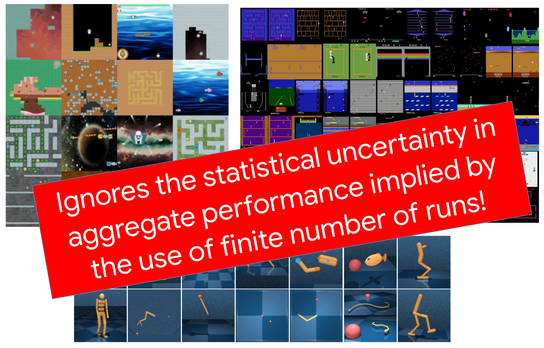
How to Avoid Fooling Oneself in Deep RL Research

Rishabh Agarwal, Max Schwarzer, Pablo Samuel Castro, Aaron Courville, Marc G. Bellemare

agarwl.github.io/rliable

Deep Reinforcement Learning at the Edge of the Statistical Precipice





How to do Deep RL Research at the Edge of the Statistical Precipice?

Performance estimates based on a finite number of runs is a random variable and should be treated as such.

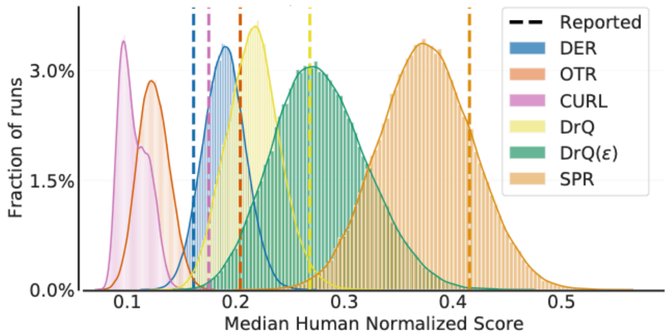
A. Use interval estimates as opposed to point estimates for accounting for uncertainty in results.

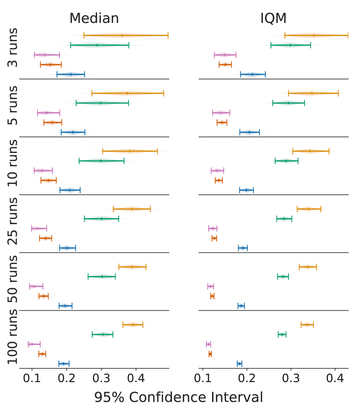
B. More is more: Performance profiles for qualitative summarization

C. Better aggregate performance measures from robust statistics such as interquartile mean (IQM) and probability of improvement.

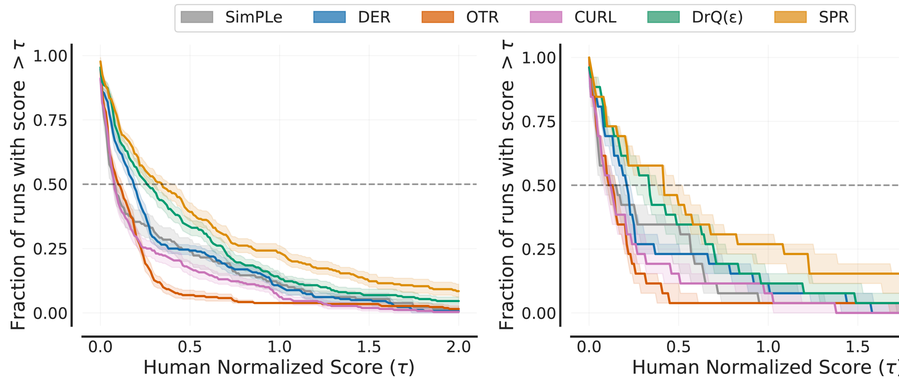
D. Provide individual runs for thorough statistical comparisons.

Report Interval Estimates rather than Point Estimates.

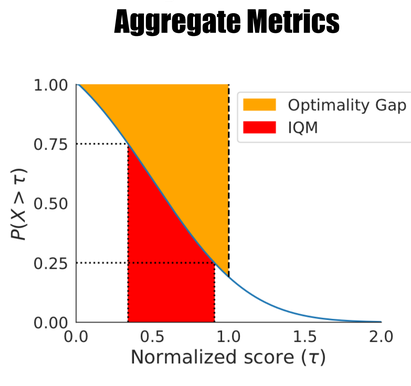




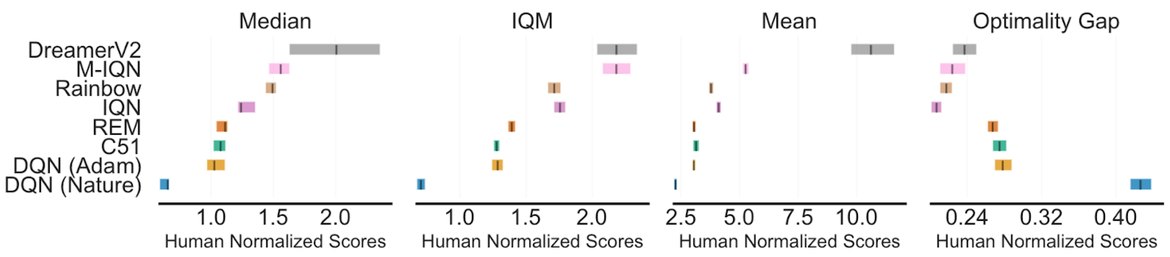
Performance Profiles for a Bird's Eye View.



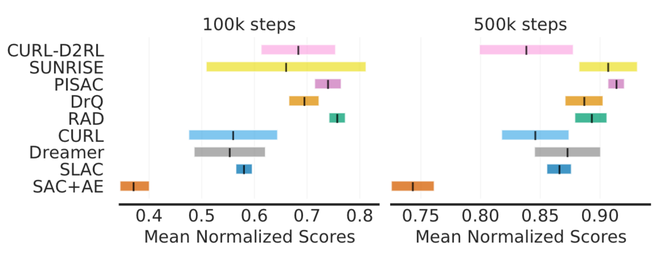
Statistically Robust and Efficient Aggregate Metrics



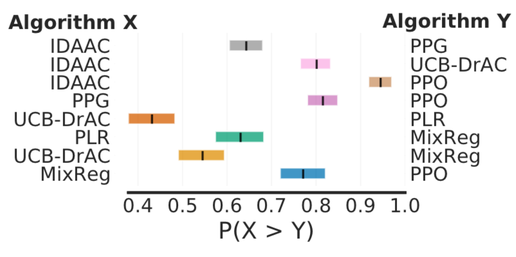
Re-evaluating Performance on Prevalent Deep RL Benchmarks



Atari 2600 Games



DM Control



ProcGen